



Services  
Canada

Services  
Canada

99-908E

c. 3

# *Census Recensement*

## *Canada 1986*

Report 104

USER'S GUIDE  
TO  
1986 CENSUS DATA  
ON  
ETHNIC ORIGIN

Canada

*Symplectic*  $\mathbb{C}P^2$ ,  $\mathcal{M}_0(\mathbb{C}P^2)$



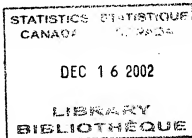
Statistics  
Canada

Statistique  
Canada

Census  
Recensement

Canada 1986

Reference



99-5C-035

**USER'S GUIDE  
TO  
1986 CENSUS DATA  
ON  
ETHNIC ORIGIN**

Published under the authority of the Minister of Industry,  
Science and Technology

All rights reserved. No part of this publication may be  
reproduced, stored in a retrieval system or transmitted in  
any form or by any means, electronic, mechanical,  
photocopying, recording or otherwise without prior written  
permission of the Minister of Supply and Services Canada.

December 1990

Price: Canada, \$23.00  
Other Countries, \$24.00

ISBN-0-660-13349-0

Ottawa

La version française est disponible sur  
demande (ISBN - 0-660-92748-9).



## ACKNOWLEDGEMENTS

This document was prepared by Pamela M. White, of the Housing, Family and Social Statistics Division, Statistics Canada.

The author would like to thank Jane Badets and Cynthia Wilson for their assistance and advice. Thanks are also extended to the staff of the Housing, Family and Social Statistics Division, in particular Harry Champion and Michael Crew.

## TABLE OF CONTENTS

	Page
I. Introduction .....	1
II. What is Ethnic or Cultural Origin? .....	2
III. Data Collection .....	6
IV. Question Response .....	7
V. Edit and Imputation .....	10
VI. Pre-Derive .....	11
VII. Imputation .....	20
VIII. Post-Derive .....	22
IX. Data Weaknesses .....	24
References .....	26
Appendix A. Ethnic Groups at LOAD Stage .....	27
Appendix B. List of Statistics Canada Publications Showing 1986 Ethnic Origin Data .....	31
Appendix C. Population and Occupied Private Dwellings, Showing Estimates for Incompletely Enumerated Indian Reserves and Indian Settlements, for Canada, Provinces and Territories, 1986 .....	32
Appendix D. Mother Tongue and Ethnic Origin Coding Errors .....	33

### List of Tables

1. Comparison of Non-Response Rates, 1981 and 1986 Ethnic Origin Data .....	7
2. Distribution of the Frequency of Double, Triple, etc. Multiple Ethnic Responses for Canada and E&I Bases, 1986 .....	8
3. Ethnic Origin, Single and Multiple Response, 1986 Census, Canada .....	9
4. Distribution Showing the Frequency of Identical Cases In ETO40I, ETO41I, ETO42I for Canada and E&I Bases, 1986 Census .....	12
5. Number of Unweighted, Unrounded Ethnic Entries Deterministically Assigned during Pre-Derive, 1986 Census, Canada .....	14
6. Number of Records Potentially at Risk for Being In Conflict, Canada, 1986 Census .....	16
7. Ethnic Origin by Processing Flag, Unweighted E&I Data, Canada, 1986 .....	21
8. Aboriginal Responses by Home Language and Mother Tongue, 1986 Census, Canada, 20% Sample Data .....	24

## I. INTRODUCTION

The 1986 Ethnic Origin User's Guide updates the 1981 Census Users Guide for Ethnic Origin written by W.O. Boxhill. The purpose of the 1986 version is to provide basic information on the 1986 Census concept of ethnic origin and to explain the data processing strategies used to edit the data.

## II. WHAT IS ETHNIC OR CULTURAL ORIGIN?

### 1. Definitions

The Users' Guide to 1981 Census Data (Boxhill:1986) explained the complexity of the concept of ethnic origin in the following ways:

There is ample evidence that ethnicity and ethnic origin are complex multidimensional concepts, definitions of which have generally been avoided (Isajiw: 1979). The Harvard Encyclopedia of American Ethnic Groups (Thernstrom, Orlov and Handlin: 1981) provides a host of possible considerations in the identification of ethnic groups and measurement of ethnicity, with combinations that may vary considerably for different groups. These are:

1. common geographic origins;
2. migratory status;
3. race;
4. language or dialect;
5. religious faith or faiths;
6. ties that transcend kinship, neighbourhood, and community boundaries;
7. shared traditions, values and symbols;
8. literature, folklore and music;
9. food preferences;
10. settlement and employment patterns;
11. special interests in regard to politics in the homeland and in the United States;
12. institutions that specifically serve and maintain the group;
13. an internal sense of distinctiveness;
14. an external perception of distinctiveness.

The authors nonetheless concede that defining ethnic groups is "no simple matter because there is as yet no consensus about the precise meaning of ethnicity" (Thernstrom, Orlov and Handlin: 1981, p. 5), and about how these 14 measures should be combined.

Boxhill mentions other important sources which bear repeating concerning the conditions and definitions of ethnic group and ethnicity, the more relevant being the definition provided by Sill: 1968, p. 167 in the International Encyclopedia of the Social Sciences:

"An ethnic group is a distinct category of the population in a larger society whose culture is usually different from its own. The members of such a group are, or feel themselves, or are thought to be, bound together by common ties of race or nationality or culture."

Other definitions cited by Boxhill include the elaboration provided in The Encyclopedia of the Social Sciences by Seligam: 1931, p. 607 which stresses the separateness experienced by many groups living within the boundaries of a larger cultural or national community.

Perhaps one of the more insightful descriptions is the one provided by Boxhill from Krauter and Davis (1978, p. 3).

"An ethnic group is not one because of the degree of measurable or observable difference from other groups; it is an ethnic group, on the contrary, because the people in it and the people out of it know that it is one; because both the ins and outs talk, feel and act as if it were a separate group."



Ringer and Lawless (1989) also stress the concept of we-ness and they-ness in the formation and maintenance of ethnic groups. In their book entitled Race, Ethnicity and Society, they note:

"Ethnic and racial groups are not completely autonomous and self-contained entities. They are, instead, part of a larger societal system that influences, shapes, and particularly in the case of racial groups may even define their very character and determine their life circumstances. As such, these groups are continually besieged by two sets of dynamic forces, which are frequently in opposition to each other: specifically, an internal set that serves to establish and to maintain the group's distinctive we-ness, and an external set that serves to shape and designate its they-ness."

## 2. 1986 Census Ethnic Origin Question

The 1986 ethnic origin question asked respondents to report their ethnic or cultural 'roots'. The question was as vague in 1986 as it was in 1981, with the exception that at least in 1981 the question contained a reference point: "on first coming to this continent". This temporal pinpointing of roots was not part of the 1986 question.

### **The 1986 Census Ethnic Origin Question and Guide**

The 1986 Census variable ethnic origin referred to the "ethnic or cultural roots" of the population as explained in the Census Guide items shown below.

#### **Question 17**

*Ethnic or cultural group refers to the "roots" or ancestral origin of the population and should not be confused with citizenship or nationality. Canadians belong to many ethnic and cultural groups, such as Inuit, North American Indian, Métis, Irish, Scottish, Ukrainian, Chinese, Japanese, East Indian (from the subcontinent of India, Dutch, English, French, etc.*

*Note that in cases where you use language as a guide to your ethnic group, you should report the specific ethnic group to which you belong, e.g., Haitian rather than French, Austrian rather than German.*

*The ethnic origin question will provide information which is used extensively by many ethnic or cultural associations in Canada to study the size, location, characteristics and other aspects of their respective groups.*

The 1986 definition differed little from the definition provided in the 1981 Census Guide:

#### **Question 26**

*Ethnic or cultural group refers to the "roots" of the population, and should not be confused with citizenship or nationality. Canadians belong to many ethnic and cultural groups — English, French, Irish, Scottish, Ukrainian, Native Indian, Chinese, Japanese, Dutch, etc.*

*If applicable in your case, a guide to your ethnic origin may be the language which you or your ancestors used on first coming to this continent, e.g., Dutch, Japanese. Note, however, that in cases where a language is used by more than one ethnic group, you should report the specific ethnic group, e.g., Austrian rather than German.*

*For Native Peoples, the phrase "on first coming to this continent" should be ignored.*

*Métis are descendants of people of mixed Indian and European ancestry who formed a distinct socio-cultural entity in the 19th century. The Métis have gone on to absorb the mixed offspring of Native Indian people and groups from all over the world.*

While the concept of ethnic or cultural "roots" did not change, the use of the phrase 'on first coming to this continent' did serve to specify a temporal point from which to trace the ancestral ethnic lineage. This reference point was dropped from the 1986 ethnic question as it was viewed as being inappropriate for persons of aboriginal backgrounds. In 1986, respondents were asked: 'To which ethnic or cultural group(s) do you or did your ancestors belong?' As well, respondents were asked to mark or specify as many groups as apply.

Several of the ethnic groups shown on the questionnaire were changed for the 1986 Census question. For example, status and non-status Indian were replaced by North American Indian. The positioning of groups on the question changed so that the groups would be shown in relative numerical order. Thus the position of Chinese and Polish altered on the 1986 question.

The mark-box Black was added to the 1986 Census question in response to the recommendations made by the Abella commission on Equity in Employment and the Parliamentary Committee report entitled Equality Now! regarding the need for data on Canada's visible minority populations.

As well, two more write-in spaces were added to the questionnaire. This allowed respondents to give up to 3 groups not already pre-specified on the question.

### 1981 Ethnic Origin Question

- Mark or specify as many as applicable

- Other ethnic or cultural group(s). For example, Portuguese, Greek, Indian (India), Pakistani, Filipino, Japanese, Vietnamese. (specify below)

Other (specify) \_\_\_\_\_Other (specify) \_\_\_\_\_Other (specify) \_\_\_\_\_

- (See Guide for further information.)*

- Other (specify) \_\_\_\_\_

### III. DATA COLLECTION

The 1986 ethnic origin question was asked of one in five Canadian households. The information collected from these households is 'weighted' to produce estimates for the total Canadian population. Such an approach reduces the overall response burden, while at the same time provides statistically reliable information on the demographic, socio-economic and cultural characteristics of Canadians.

In 1986, more than 98% of Canadians were enumerated via the self-enumeration method. Prior to Census Day, June 3, 1986, questionnaires were delivered for completion by household members on this date. Completed documents were mailed back to Statistics Canada or picked up by census representatives. The canvasser method was used for less than 2% of the population, with census staff visiting homes to collect the required information through interview. Canvasser enumeration was undertaken mainly in remote or northern areas and on Indian reserves where irregular mail service makes mail-back impractical and long distances rule out the pick-up method.

#### IV. QUESTION RESPONSE

##### 1. Non-response

In 1986, the non-response rate of 2.5% for the ethnic origin question was slightly higher compared with the 1981 level of 2.3%. Table 1 shows the regional variations in non-response. Non-response was lowest in Quebec and the Maritimes and highest in the West and among the 2C (overseas) populations.

TABLE 1: Comparison of Non-Response Rates, 1981 and 1986 Ethnic Origin Data

E & I Bases	1986 Data		1981 Data	
	N	%	N	%
Canada	131,330	2.50%	116,960	2.30%
East	11,010	2.38%	9,680	2.20%
Quebec	28,215	2.12%	24,980	1.90%
Ontario	44,925	2.48%	40,630	2.40%
West	46,195	2.86%	40,630	2.70%
2C	990	3.89%	910	4.00%

Source: Statistics Canada, unpublished data, unweighted, data at LOAD stage.

There was no mandatory follow-up of non-response for the 1981 and the 1986 ethnic origin questions. As an ethnic origin response was required from every person in Canada, non-response was an unacceptable situation which had to be resolved through the assignment deterministically or imputation of responses. This occurred in a process called Edit and Imputation, which is described in Section V.

##### 2. 1986 Response Patterns -- Single and Multiple Responses

In 1981, multiple ethnic origin responses were not encouraged, though they were accepted. Changes in 1986, such as adding the instruction to mark or specify as many [groups] as apply and the addition of two write-in spaces (three spaces in all for 1986) compared with the one space provided in 1981, increased the level of multiple ethnic origin response from 11% in 1981 to 28% in 1986.

In 1986, at the Load stage of data processing, 2.5% of respondents gave no ethnic origin, 70.4% gave one group, 16.7% gave 2 groups, 6.6% reported 3 groups and 3.8% had an ethnic background that included 4 or more groups. As shown in Table 2, residents of Quebec (3.8%) were much less likely to give 2 or more groups compared with respondents living in Ontario (12.9%) or the West (14.6%).

A single ethnic response occurred when only one mark box answer was reported or only one write-in response was specified by the respondent. As shown on Table 3, of the 3,694,380 respondents who gave one ethnic group, 90% were made in the mark-in boxes (3,324,515) and 10% were single response write-in answers (369,810).

French was the most common single mark-box response (1,224,480), followed by English (934,350), German (185,160), Scottish (170,230) and North American Indian (150,930). As indicated in Appendix A, the most frequent write-in response were the East Indian (46,403), Portuguese (43,839) and Norwegian (42,867) groups.

In total, mark box answers accounted for 85% of respondent's reported responses. About 7.3% of respondents used only the write-in spaces to report their ethnic origin(s). A further 5.2% responded using both write-ins and mark boxes and as previously noted, 2.5% of respondents did not answer the ethnic origin question.

TABLE 2: Distribution of the Frequency of Double, Triple, etc. Multiple Ethnic Responses for Canada and E & I Bases, 1986

REPORTED ETHNIC RESPONSE	CANADA N	%	EAST N	QUEBEC N	ONTARIO N	WEST N	2C N
NON RESPONSE	131,330	2.50%	11,010	28,215	44,925	46,195	990
SINGLE RESPONSE	3,694,390	70.44%	317,590	1,209,810	1,170,270	982,150	14,560
MULTIPLE RESPONSES	1,418,720	27.05%	134,505	90,575	595,930	587,810	9,905
2	876,365	16.71%	87,725	68,265	362,870	351,800	5,710
3	344,830	6.58%	31,680	15,355	152,050	143,100	2,640
4	140,160	2.67%	11,555	5,035	58,990	63,435	1,145
5	41,615	0.79%	2,705	1,385	16,540	20,685	305
6	11,350	0.22%	650	365	4,050	6,205	80
7	2,955	0.06%	130	120	1,000	1,695	20
8	845	0.02%	25	25	275	515	5
9	255	0.00%	10	10	55	180	0
10	90	0.00%	15	5	25	50	0
11	45	0.00%	5	0	10	30	0
12	30	0.00%	5	5	10	10	0
13	5	0.00%	0	0	0	5	0
14	15	0.00%	0	0	5	10	0
15	140	0.00%	5	5	55	80	0
16	20	0.00%	0	0	10	10	0
17	5	0.00%	0	0	5	0	0
18	10	0.00%	0	0	0	10	0
TOTAL POPULATION	5,244,430	100.00%	463,100	1,328,600	1,811,125	1,616,150	25,450

Source: Statistics Canada, unpublished data, unedited and unweighted, LOAD stage.

Due to rounding, columns may not always equal the total.

Table 3: Ethnic Origin, Single and Multiple Response, 1986 Census, Canada

	Count	Percentage
Non-Response	131,330	2.50%
Single Response	3,694,380	70.44%
Single Mark-in	3,324,515	63.39%
French	1,224,480	23.35%
English	934,350	17.82%
Irish	138,590	2.64%
Scottish	170,230	3.25%
German	185,160	3.53%
Italian	135,700	2.59%
Ukrainian	82,910	1.58%
Dutch	69,670	1.33%
Chinese	68,435	1.30%
Jewish	46,820	0.89%
Polish	43,325	0.83%
Black	32,150	0.61%
Inuit	23,465	0.45%
North American Indian	150,930	2.88%
Metis	18,310	0.35%
Single Write-in	369,810	7.05%
ETO40I	368,685	7.03%
ETO41I	865	0.02%
ETO42I	260	-
Multiple Response	1,418,720	27.05%
Multiple Mark-ins	1,124,835	21.45%
Multiple Write-ins	18,085	0.34%
ETO40I + ETO41I	17,965	0.34%
ETO40I + ETO42I	45	-
ETO41I + ETO42I	75	-
Multiple Mark-ins and Write-ins	246,215	4.69%
Mark-ins + ETO40I	242,800	4.63%
Mark-ins + ETO41I	2,900	0.06%
Mark-ins + ETO42I	515	0.01%
Other Multiple Mark-ins and Write-ins	29,585	0.56%
Total	5,244,430	100.00%

Source: Statistics Canada, unpublished data, unweighted and unedited, LOAD stage.

Due to rounding, percentages may not always equal 100.

## **V. EDIT AND IMPUTATION**

All census information was validated via a processing routine referred to as 'edit and imputation'. During this process, responses were deemed to be valid or in conflict on the basis of predetermined edit rules. Valid data were accepted without modification. Conflicts were resolved following a rule of minimum change to respondent provided data. In the case where no ethnic origin was provided by the respondent, responses were imputed according to predetermined imputation rules.

The processing of the ethnic origin variable was accomplished in three stages: pre-derive, imputation and post-derive.



## VI. PRE-DERIVE

The purpose of the pre-derive editing phase of the 1986 ethnic origin variable was to remove duplicate write-in entries, resolve the South Asian/Aboriginal conflict, simplify retrieval of responses for the Black groups and to remove religious type responses. As well, where some children in a census family lacked a response, responses were provided based on predetermined edit rules.

### 1. Duplicate write-in responses

One of the first changes made to the responses was the removal of duplicate write-in responses. The 1986 ethnic origin question contained three (3) write-in spaces. In order to ensure that unique responses were given for each respondent, the write-ins were checked so that there would be no duplication of response. As well, the first write-in entry was always to be found in the first write-in space. Table 4 shows that duplicate answers made-up about .03% of all write-in responses.

### 2. Re-coding of entries

After the removal of duplicate entries, some coded entries were deterministically re-assigned. At the manual coding stage, codes were assigned to groups not intended to be carried to the retrieval data base.

For example, as the 1986 Census did not feature a religion question, it was expected that some respondents would choose to give a religious group rather than an ethnic one. It was decided to code these answers and in the cases of low counts reassign these religious groups to ethnic ones.

Table 5 shows the frequency of the groups deterministically changed during the pre-derive stage. The changes are described below.

#### a) V555-Other-Nie

During the coding operation of Regional Office Processing (ROP), the code V555-Other-Nie was assigned in the Newfoundland Regional Office Processing centre for the write-in entries of 'Settler', 'Native Settler' and 'Kablunuk'. During the pre-derive, the write-in entry was made A\_Blank and ETO37I-Inuit was made true. Outside of Newfoundland, the entry of V555-Other-Nie was treated as an Invalid code that could have been caused by incorrect coding or key entry error.

In Labrador, persons of mixed Inuit and non-Inuit heritage are often referred to as Settlers. This is a unique term which applies just to the Labrador region.

#### b) D016-Amerindian-Nie

In 1986, 31,834 respondents gave a tribal or Indian nation type of write-in response. As individual tribal or Indian nation responses could not be identified separately, they were given the code of 016. During the pre-derive, the write-in was made A\_BLANK and the mark box ETO39I-North American Indian was made TRUE.

**TABLE 4: Distribution Showing the Frequency of Identical Cases in ETO40I, ETO41I, ETO42I for Canada and  
and E & I Bases, 1986 Census**

# OF CASES WHERE	CANADA		EAST	QUEBEC	ONTARIO	WEST	2C
	N	%	N	N	N	N	N
ETO40=ETO41/ETO42(*)	1,075	0.02%	15	100	525	430	5
ETO40=ETO42/ETO41(*)	55	0.00%	0	5	30	15	0
ETO41=ETO42/ETO40(*)	50	0.00%	0	0	30	15	0
ETO40=ETO41=ETO42(*)	135	0.00%	0	20	75	40	0
ELSE	5,243,115	99.97%	463,085	1,328,480	1,810,460	1,615,650	25,445
<b>TOTAL</b>	<b>5,244,430</b>	<b>100.00%</b>	<b>463,100</b>	<b>1,328,600</b>	<b>1,811,125</b>	<b>1,616,150</b>	<b>25,450</b>

(\*) - Cases where ETO40=ETO41=ETO42=A\_BLANK are excluded.

SOURCE: Statistics Canada, unpublished data, unedited and unweighted data, LOAD stage. Due to rounding, percentages may not always equal 100.

c) Removal of Ethno-Religious Responses

In 1986, a census religion question was not asked. In expectation of religious type response to the ethnic origin question, codes were assigned during regional office processing. In the pre-derive stage of processing, these religious type entries were re-assigned to certain ethnic groups. The following changes were made:

- i) The write-in of Hindu was made East Indian, n.i.e.
- ii) The write-in of Muslim-Islamic was assigned to E. Indian if the mother tongue was of the South Asian type. For all other mother tongues, the responses were made Arab, n.i.e. Assignment at the family level assured that children in the family who may have English and French as a mother tongue maintained the responses of their parents.
- iii) The write-in of Hutterite was made German.
- iv) The write-in of Mennonite was made German.
- v) The write-in of Buddhist was made Other Asian, n.i.e.
- vi) The write-in of Sikh was made Punjabi.
- vii) The group Other Religious groups was made Other, n.i.e.

d) Cypriote Responses

When the coding system for 1986 was planned, it was expected that respondents would report Turkish Cypriote, Greek Cypriote and Cypriote. However, response levels for these groups were too low to take the data to the retrieval data base. Therefore the following changes were made:

- i) Turkish Cypriote was made a multiple response -- Turk and Cypriote.
- ii) Greek Cypriote was made a multiple response -- Greek and Cypriote.
- iii) Cypriote was maintained.

e) Black Responses

In order to improve the ease of retrieval the following changes were made to the Black entries:

- i) The write-in of West Indian Black changed so that the Black mark-in box ETO36I was made TRUE and the write-in changed to West Indian.
- ii) The write-in of Canadian Black was changed so that the Black mark-in box ETO36I was made TRUE and the write-in changed to Canadian.
- iii) The write-in of American Black was changed so that the Black mark-in box ETO36I was made TRUE and the write-in changed to American.

**TABLE 5: Number of Unweighted, Unrounded Ethnic Entries Deterministically Assigned during Pre-Derive, 1986 Census, Canada**

Ethno-Religious Groups	ETO40I	ETO41I	ETO42I	Total
T111- Buddhist	25	0	0	25
T112- Hindu	920	120	10	1,050
T113- Sikh	910	255	35	1,195
T114- Muslim-Islamic	715	210	40	965
T116- Doukhobor-Freedomite	280	60	5	340
T117- Hutterite	1,325	50	80	1,455
T118- Mennonite	4,590	390	25	5,005
D016- Amerindian, nie	29,120	2,345	365	31,835
U123- Greek Cypriot	65	0	0	65
U124- Turk Cypriot	0	0	0	0
U555- Other, nie (Inuit Settler)	285	5	0	290
T119- Other Religious Groups	115	25	10	150
I027- West Indian Black	65	0	5	70
I028- Black American	15	0	0	15
I029- Canadian Black	95	10	0	105
I030- Black, nie	60	5	5	70

Source: Statistics Canada, unpublished data, unweighted and unedited, LOAD stage.

- iv) The write-in of Black, nie was changed so that the Black mark-in box ETO36I was made TRUE and the write-in was made A,BLANK.

In order to remove any conflict between the write-in response of African Black, Other African and the Black mark-in box, the following conflict resolutions were made.

In the case of a respondent having Other African and the Black mark-in box marked, the Other African was changed to African Black. Other African is supposed to represent the non-Black African groups.

### 3. Resolution of Aboriginal and South Asian Response Conflicts

Boxhill (1985) discovered that during the 1981 Census processing a response error had been permitted to slip through undetected. Over 2,000 aboriginals had Asian mother tongues and a non-Christian religion such as Sikh or Hindu and were born outside of Canada or the United States, for example, India, Pakistan, Fiji, Tanzania, Guyana or Sri Lanka. For 1986, it was decided not allow such conflicts to be remain unedited.

The conflict between the Aboriginal and South Asian responses occurs when respondents write-in the group Indian or when a person of South Asian origin checks the box North American Indian. In 1981, the write-in of 'Indian' was coded to Amerindian, n.i.e. In 1986, the code for Indian was not shown in the Population Code Book. When Indian was written in by the respondent, coders were required to examine the mother tongue and place of birth of the respondent before assigning a code. This procedure reduced the possibility of confusion between a South Asian response and an aboriginal one, but added to the work load of the coder. Coders are expected to code a required number of EA boxes per day and such additional steps slowed down the coding process. Thus the procedure was not always followed.

Another respondent response problem was also noted. It is quite common for Haitians and other francophones from the Caribbean who are of mixed racial backgrounds to call themselves Métis. In the true sense of the word, they are Métis. However, they are not of mixed North American Indian background, though they could be of Arawak or of some other indigenous Caribbean Indian background. It is expected that as Haitian immigration increases and persons of mixed backgrounds become more prevalent, the term Métis will no longer be reserved exclusively in the minds of the respondent for persons of mixed Aboriginal and non-aboriginal backgrounds. The 1986 editing strategy also took this type of respondent pattern into consideration.

When the 1986 Edit and Imputation specifications were originally designed it was assumed that the 2A question #7 could be used to assign non-response and to validate the responses on respondents who claimed to be of aboriginal origin. Unfortunately the 2A question #7 did not work (Hagey:1987) and the data could not be used in the edit and imputation of the 2B ethnic responses.

The failure of Question #7 raised the specter of doubt for the aboriginal responses obtained from the 2B ethnic origin question. It was decided to submit all the data to a very close verification procedure. This would eliminate the South Asian/Aboriginal conflicts and also ferret out the erroneous responses provided by respondents who did not understand the term Inuit, which had been a major problem with the Question #7.

One consequence of the failure of question #7 was to alter the edit and imputation strategy for the 1986 ethnic origin variable. As the editing structure for processing of the ethnic variable had already been established certain compromises had to be made. This resulted in all records going through the Aboriginal/South Asian data clean-up. Unfortunately, it was not possible to accurately monitor the specific changes which occurred as all records were judged to be GOOD, BAD or SUSPECT vis-a-vis the Aboriginal/South Asian clean-up and assigned the flag. This flag could later be written over, should there be a problem with other elements of the response pattern. Table 6 shows the entries at the LOAD stage which were under the most intense inspection.

TABLE 6: Number of Records Potentially at Risk for Being in Conflict, Canada, 1986 Census

ETO37- Inuit	26,985			
ETO38- North American Indian	235,240			
ETO39- Metis	41,250			
	ETO40I	ETO41I	ETO42I	Total
D016- Amerindian, nie	29,120	2,345	365	31,835
V555- Other, nie	285	5	0	290
Q086- Bengali	100	10	5	115
Q087- Gujarati	175	55	5	230
Q088- Punjabi	1,490	300	10	1,795
Q089- Tamil	295	120	5	420
Q090- East Indian, nie	46,405	1,725	205	48,330
Q091- Bangladeshi	290	20	0	310
Q092- Pakistani	5,430	510	40	5,985
Q093- Sri Lankan	1,290	60	10	1,360
Q094- Singhalese	195	10	0	210
T112- Hindu	920	120	10	1,050
T113- Sikh	910	255	35	1,195
T114- Muslim Islamic	720	210	40	965

Source: Statistics Canada, unpublished data, unweighted and unedited, LOAD stage.

It is important to note, that at this stage of the processing there was no summary variable. Thus, it is responses which are under consideration not the full string of groups provided for the respondent. This is one of the difficulties in processing a variable with three write-ins, as the action in any of the three write-ins or 15 mark-in boxes can trigger a change to the Flag value.

There were several stages of evaluation which all records went through. It should be noted that no changes could be made to the mother tongue data as this 2A variable had already been processed. Therefore if there was a conflict between an ethnic response and mother tongue, for example the mother tongue was Hindi and the ethnic origin was Inuit, the ethnic origin entry would have to be changed.

All records were inspected against the mother tongue data. Suspect records which were in potential conflict with mother tongue were then evaluated against residence type, (on Indian reserve or settlement), and finally against place of birth and citizenship. Also note that the 2B place of birth and citizenship had not been processed at this stage.

a) Mother Tongue

The checking of the records went as follows. If an aboriginal entry (Inuit, North American Indian, Métis) had a mother tongue of E01\_CREE to E028\_WAKASHAN, E030\_INUKTITUT, or E031\_MOHAWK, the record was considered GOOD. The aboriginal origins were retained and no further processing was required at this time.

If the aboriginal ethnic origin respondent had a non-official language other than an aboriginal one, the ethnic value was considered to be a BAD one and the aboriginal value was removed. This would occur for example if the mark-box ETO391\_NORTH AMERICAN INDIAN was marked true and the mother tongue was Hindi. As the mother tongue value could not be changed, the ethnic had to be changed. The mark-box was made FALSE and no further processing occurred for the BAD records at this time.

If the aboriginal record had a mother tongue of English, French, a multiple English and French or a multiple response with English and/or French with a non-official language which was not an aboriginal one, then the record was considered SUSPECT. As well, if the record had a mother tongue of E029\_AMERIND\_NIE, the aboriginal record was considered SUSPECT. During coding a large number of South Asians were coded to the Amerindian, nie language code. All suspect entries required further processing.

The exception to this was the mother tongue E029\_AMERINDIAN\_NIE. In this case, it was the mother tongue response which was incorrect and not the ethnic value. Given that this mother tongue value had been identified at release as being polluted, it was decided to maintain the remainder of the 2B characteristics for these records.

Thus after the mother tongue check of the Aboriginal and South Asian ethnic responses, the GOOD and the BAD records had been established. Only the SUSPECT records remained to be further processed.

It should be noted that the editing was done largely to the aboriginal responses with erroneous mother tongues. After processing, it became apparent that there were other data quality issues which should have also been considered (Appendix D). In particular, the aboriginal mother tongue data having odd ethnic responses. These were introduced during coding and had not been expected when the editing strategy had been conceived.

b) SUSPECT Records

i) Residence type

The SUSPECT mother tongue entries were examined by residence type. It should be noted that at this stage of processing only those aboriginal records which had an English, French, and English and French mother tongue

were examined. If these type of records were resident on a reserve or Indian settlement, they were no longer considered SUSPECT.

The records still considered SUSPECT were examined for place of birth and citizenship. The records now considered SUSPECT had an English, French or English and French mother tongue and the CSD Type was off-reserve or other than an Indian settlement.

ii) Place of Birth and Citizenship

The remaining SUSPECT records were examined by place of birth. Aboriginal respondents were expected to have a place of birth of Canada and have a citizenship of Canada, by birth, though some exceptions were allowed.

The permitted exceptions to the Canada by birth citizenship and place of birth of Canada were for aboriginal respondents with a citizenship of Canada by birth and a place of birth other than Canada, the United States or Greenland. In most cases, this would be respondents who had a place of birth of West Germany (parents in the Canadian Armed Forces).

If the SUSPECT records showed the respondent was an immigrant (citizenship other than Canada, by birth and a year of immigration), only a certain set of places of birth could be allowed -- the United States and Greenland. For all other places of birth, the SUSPECT aboriginal record would be considered BAD and the aboriginal ethnic response deleted. This would have the effect of deleting the erroneous aboriginal records from respondents with the following profile: English mother tongue, immigrant, citizenship of Canada by naturalization and place of birth of Sri Lanka and ethnic origin North American Indian.

This clean-up was not perfect but it was more extensive compared with 1981. It also ensured that the aboriginal data did not have any obvious defects. As the processing was designed to loop through the family, it would eliminate the aboriginal responses for the family which had parents with mother tongue of Hindi but the children born in Canada, non-immigrant and with an English mother tongue would also have the North American Indian value deleted.

4. Aboriginal Assignment

The final stage of the aboriginal clean-up was to deterministically assign responses to completely blank records which met the following conditions. If the ethnic question was blank but the mother tongue was Inuktitut, then ETO37U\_INUIT was made TRUE.

For respondents living on Indian reserves with an aboriginal mother tongue or an English and/or French mother tongue, and the ethnic value was blank, then ETO39U\_NORTH AMERICAN INDIAN was made TRUE. This was done to avoid the risk of using the non-aboriginal teacher or priest living on reserve as the donor record during the HOT DECK phase of imputation.

5. Family Assignment

The final phase of the pre-derive was to assign responses to children living in census families where there was a response for at least one child in the family. From time to time the person completing the census form forgets to fill in all of the questions for the younger children in the family.



As ethnic origin is an ascribed characteristic, it was decided to give the children in the family the origin of the next older child having a response. This allows for differences in mixed or blended families. For example, in a blended family the oldest child could have had an origin of German and Irish and the second child could have German and Chinese with the third child in the family having a blank ethnic response. This third child would receive the origin of next oldest child, which in this case would be German and Chinese.

#### 6. Result of Pre-derive

After the pre-derive stage of processing, the following types of records would remain.

- i) There would be a group of GOOD records for which there were answers with no conflicts (Aboriginal/South Asian) and all of the members of the family would have ethnic origin responses.
- ii) A second group would include the Blank response husband and/or a Blank response wife and the Blank response lone parent in a census family with children fully responded.
- iii) The third group would be totally Blank response records where all members of the census family would be blank. Included in this third type would be Blank non-census family members.
- iv) As well, all duplication of response write-ins had been removed. The write-ins had been sorted so that the first write-in was always ETO40U, the second was always ETO41U and the third always occurred in the ETO42U space. In the case of GOOD records, V\_BLANK was placed in the empty write-in spaces. This was done so that during imputation the record would not be judged non-response because there were write-in spaces not having an ethnic origin.
- v) Also, the mark-ins were also transferred to the variable ETMARK. ETMARK summarized in a numerical way all of the possible combinations of mark-ins.

## VII. IMPUTATION

The purpose of imputation in census processing is to find for a Blank (non-response) record a donor which is as similar as possible to the characteristics of the recipient. This was accomplished in a deterministic way at the pre-derive stage when Blank children were given the responses of their siblings.

In order to give ethnic origin response to Blank response adults in census families (husband, wives and lone parents) and non-census family persons, a response was imputed by the computer using the HOT DECK method. Conditions influencing the imputation of ethnic origins were language (mother tongue) of the donor and the recipient, CSD type, age and sex.

The mandatory matching conditions was mother tongue. For example, respondent with a Blank ethnic origin and a Greek mother tongue must match with a donor who also has a Greek mother tongue.

The non-mandatory matching conditions were sex, age and CSD type. These conditions could be relaxed within the mother tongue mandatory matching conditions if a donor could not be readily located.

The search length for a suitable donor was 50 records across a processing base (East, Quebec, Ontario, West and 2C). If a donor could not be found, then a default ethnic origin response(s) was assigned.

The 1986 imputation strategy followed the intent of the processing strategy used in 1981. However, unlike 1981, the 1986 approach did not pre-define the acceptable donor ethnic origin responses within the mother tongue mandatory constraint.

Unlike 1981, multiple ethnic origins were imputed for 1986. The SPIDER processing system works on the principle of least change. Therefore, it will look for donor record carrying a single ethnic response as this type of response requires that SPIDER transfer only one field from one record to another. This technical problem was overcome for 1986.

Table 7 shows the number of records which were given responses through the HOT DECK and by DEFAULT in 1986. About 5 records went to DEFAULT in 1986 and this occurred in the Northwest Territories.

Table 7: Ethnic Origin by Processing Flag, Unweighted E & I Data, Canada, 1986

		%
Total Population	5,244,430	100.0% *
No Imputation	4,163,185	79.4%
Pre-derive change(1)	9,455	0.2%
Aboriginal/South		
Asian Inspection(2)	939,520	17.9%
Aboriginal Assignment(3)	17,015	0.3%
Family Assignment pre-derive	15,390	0.3%
Family Assignment post-derive	22,270	0.4%
Default	5	0.0%
Hot Deck	77,595	1.5%

\* Due to rounding, totals do not equal 100%

(1) Total number of records changed because of pseudo-religious codes, black clean-up, duplication and sorting.

(2) Total number of records inspected for possible aboriginal/South Asian conflict.

(3) Total number of records assigned an aboriginal response due to non-response on reserve. Partial response reserves were carried through edit and imputation but not carried to the retrieval data base, which explains the high assignment levels on reserve.

Source: Statistics Canada unpublished E&I data.

## VIII. POST-DERIVE

After the pre-derive and imputation stages of processing two types of response pattern remained.

- i) Census families and non-family persons with complete responses.
- ii) Census families in which all of the children were non-response.

It was the function of the Post-derive operation to obtain responses for the non-response children.

The second function of the post-derive involved a final check on the aboriginal and South Asian responses to ensure that conflicts had not been inadvertently assigned or imputed.

The third task of the post-derive was to create the retrieval data base variables of ETOSING, ETO123 and ABRES.

### 1. Assignment of Parents' Responses to Blank Children

In the pre-derive, the origins of the elder siblings had been assigned to non-response children within the family. Thus at the post-derive stage of processing, no Census family existed with partial response children. On the assumption that the child's ethnic origins should be the sum of their parents, it was reasonably straight forward to assign the parents origins to the non-response children in the census family.

Clearly, this approach is not correct in a blended family or where adoption has occurred. As there was no way of determining whether the child was actually the birth child of the parents with whom he lived, the assumption was made to treat the non-response children in census families as the natural progeny of their parents.

Assignment was carried out in the following fashion:

- a) All of the mark-box entries of both the father and mother were assigned to the children.
- b) In the case of lone-parent families, the children received all of the origins of the sole parent with whom they lived. This included mark-boxes and write-in entries.
- c) In the case of two-parent families, the transfer of write-ins depended on the number of identical write-ins.

#### 3 Identical write-ins

In the case of both parents having identical write-ins, the write-ins of one parent were transferred to the children.

#### 2 Identical write-ins

In the case of two identical write-ins, the two entries would be transferred. The third write-in would be randomly selected from one parent.

#### 1 Identical write-in

In the case of one identical write-in, it would be assigned. For the remaining write-in spaces, a parent's first non-identical write-in would be selected and for the third entry the second parent would be used to select the next non-identical group.

#### No Identical Write-ins

In the case of no identical origins, a parent would be selected for the first write-in. The first write-in of the second parent would be used to form the second write-in of the child. The second write-in of the first parent would be used for the child's third write-in. Adjustments were made depending on the number of write-ins available for each parent.

#### 2. Final Check of Aboriginal and South Asian Origins

The Aboriginal and South Asian responses were checked so that any conflicts created through imputation or assignment would be removed. The mother tongue of the aboriginal responses were checked and only those records with mother tongues of English, French, English and French or an aboriginal language were permitted. The data were also checked so that no record could have both Aboriginal and South Asian origins.

#### 3. Retrieval Variables

The final phase of the processing was to create three retrieval variables: ETO123, ETOSING and ABRES.

##### a) ETO123

This variable identified the response pattern of each respondent's answer and grouped it into the category of single response or multiple response. This variable was required so that researchers would be able to combine origins yet at the same time count respondents as opposed to responses.

##### b) ETOSING

The derived variable, ETOSING showed the single responses for all of the 105 census ethnic groups. It also grouped together the multiple responses on the basis of British, French and Other origins.

This variable is useful as it permits users to show ethnic origins for geographical areas and allows for comparisons between areas without danger of double counting respondents.

##### c) ABRES

The retrieval variable ABRES identified aboriginal respondents. The variable was used by other subject matter areas in the processing of the census variables.

#### 4. 1986 Data Base

It is possible to use the ETOSING, ETO123 and the combination of the 15 logical and three write-in variables to obtain data on ethnic groups in Canada. Appendix B lists the 1986 Census publications containing ethnic origin data.

## IX. DATA WEAKNESSES

### 1. Aboriginal Data

The 1986 processing of the ethnic origin aboriginal data used the mother tongue data to check aboriginal responses. Some of the variables processed after the ethnic origin variable used the ABRES variable to check or to stratify responses. For example, the Immigration and Place of Birth processing used the ABRES variable to verify the place of birth of aboriginal respondents and to maintain the constraints used in the ethnic origin processing. The processing of income data also used the ABRES variable.

A similar strategy was not employed in the processing of the home language variable. A cross-tabulation of ethnic origin by home language shows inconsistent home languages for single response aboriginals. Table 8 uses weighted 1986 Census data to show this anomaly for the aboriginal origins by home language.

**TABLE 8: Aboriginal Responses by Home Language and Mother Tongue,  
1986 Census, Canada, 20% Sample Data**

	Home Language	Mother Tongue
Total Population	711,720	711,720
Single Responses	655,210	669,745
English	514,215	480,875
French	47,560	56,555
Non-official Language	93,435	132,315
Aboriginal	93,195	132,320
Spanish	15	-
Other European	155	-
Asiatic	65	-
Multiple Responses	56,515	41,975

Sources: Profile of Visible Minorities and Aboriginal Peoples, 1986 Census,  
20% Sample, January 1990.  
Retrieval Data Base 2B ethnic origin data.

Another weakness in the ethnic strategy was the emphasis on the ethnic aboriginal responses as being the source of error. Appendix D shows that through poor coding of either the mother tongue or the ethnic data, errors were introduced to the aboriginal data. A resolution of this problem in 1986 would have been to have removed all ethnic responses other than aboriginal ones from respondents having aboriginal mother tongues. This would have been a drastic step, but it would have eliminated the obvious errors in the ethnic response for several mother tongues.

Unfortunately, it is not always the case that the mother tongue is correct and the ethnic response wrong. However, as mother tongue is processed first, it is not possible to change an incorrectly coded or imputed mother tongue value. Also, as mother tongue is a 2A variable, no cross-editing between 2A and 2B data sets was permitted in 1986. Thus, decisions were made to accept or reject a mother tongue response without consideration of the 2B characteristics which might contradict the imputed or assigned mother tongue response.

It is important on the other hand not to lose perspective as the margin of error was often very small and frequently involved less than 50 respondents. For example, 50 weighted respondents corresponds to about 10 respondents who received a 2B questionnaire. So in the case of imputation or coding errors, in the order of 2 to 3 households were coded incorrectly or imputed oddly.

Another issue also occurred with the aboriginal data in 1986. While the 1986 aboriginal data do not have the data deficiencies evident in 1981, there was a coverage problem. An estimated 45,000 residents of 136 Indian Bands and Settlements refused to participate in the 1986 Census. This is a worrisome situation as the 1986 experience represents a marked increase from the 5,000 aboriginals who refused to participate in 1981.

Appendix C lists the reserves and gives estimates for the missed populations.

## 2. Coding Errors

As already discussed, response error was introduced into the mother tongue, home language and ethnic origin data during the manual coding phase. Problems were noted during limited field observation sessions by Subject Matter officers. The full extent of the coding errors were not evident until processing was complete.

It was observed that manual coders sometimes used the incorrect sections of Population Code Books to code the language and ethnic variables. This resulted in the ethnic codes being used on the language entries and the language codes being used to code ethnic groups.

Another source of error occurred when coders took liberty with spelling and confused responses. For example, the write-in 'slavey' received the code for 'slavic'. This error was noticed by Subject Matter officers and resulted in EA boxes being re-coded.

Before release of the 1986 ethnic origin data, it was discovered that two ethnic groups had been so poorly coded that they were deleted from the retrieval data base. A check of the groups Franco-Manitoban and Franco-Ontarian by mother tongue showed the two ethnic groups had very low levels of French mother tongue and high English, Swedish and Norwegian mother tongues. This was not what had been expected.

Further inspection of the actual questionnaires showed that respondents had actually reported a Swedish or Norwegian ethnic origin. However, the coders had used the mother tongue codes for Swedish and Norwegian to code the ethnic entries. Unfortunately, the mother tongue code numbers corresponded to the ethnic origin codes for Franco-Manitoban and Franco-Ontarian. It was decided to merge these two ethnic codes with the Swedish and Norwegian ethnic groups and to drop the Franco-Ontarian and Franco-Manitoban retrieval values.

The Franco-Manitoban and Franco-Ontarian situation were the cases with the most obvious coding errors. Users of the ethnic origin data, especially for small geographic areas, may also notice the following anomalies:

- i) Arab entries which were coded to Peruvian
- ii) Czech entries which were coded to Puerto Rican.

In 1991, automated coding will replace manual coding of the cultural variables. This should reduce the overall level of error introduced by the coding operation. It is expected that the quality of the cultural data will show improvement, especially at the small area, as coding oddities as evident in Appendix D should not be repeated in 1991.

## REFERENCES

Boxhill, W.O., 1981 Census Data on the Native Peoples of Canada, Canadian Statistical Review. February, 1985, pp. vi-xix.

Boxhill, W.O. 1981 Census Guide to Ethnic Origin Data. Catalogue 99-953, 1986.

Hagey, J., Data Quality Notes on Question 7 (Aboriginal Status), Statistics Canada, July 1987.



## **APPENDIX A**

### **ETHNIC GROUPS AT LOAD STAGE**

ETO25 to ETO39 correspond to the mark box entries shown on the 1986 Ethnic Origin question.

ETO40, ETO41, ETO42 refer to the three write-in spaces provided on the 1986 Ethnic Origin question.

ETO40I, for example, refer to the write-ins coded into the first space and the data are at the LOAD or pre-edit stage.

Table 1: Distribution of Ethnic Origins, ETO401, ETO411, ETO421, 1986 Census Data, Unweighted, Unrounded, E&amp;I, 2B Merge

ORIGIN	ETO401	ETO411	ETO421	TOTAL WRITE-INS	% ETO401 OF TOTAL	% ETO411 OF TOTAL	% ETO421 OF TOTAL
TOTAL	5,244,429	5,244,429	5,244,429	---	---	---	---
A_BLANK	4,580,976	5,193,014	5,238,399	---	---	---	---
B004_BRITISH_NIE	425	69	27	522	81.42%	13.22%	5.17%
B005_OT_BRITISH	907	137	14	1,058	85.73%	12.95%	1.32%
B008_WELSH	27,276	2,391	316	29,984	90.97%	7.97%	1.05%
C008_ACADIAN	1,645	37	6	1,690	97.34%	2.19%	0.36%
C009_FRANCO_MAN	566	62	4	632	89.56%	9.81%	0.63%
C010_FRANCO_ONT	518	96	5	619	83.68%	15.51%	0.81%
C011_QUÉBÉCOIS	748	44	3	796	93.97%	5.53%	0.38%
C012_FRENCH_CAN	445	50	7	502	88.65%	9.96%	1.39%
D016_AMERIND_NIE	29,118	2,343	363	31,834	91.47%	7.36%	1.14%
E017_CANADIAN	22,118	1,295	241	23,656	93.49%	5.47%	1.02%
F018_AMERICAN	3,349	1,164	217	4,732	70.77%	24.60%	4.59%
G070_ARGENTINIAN	434	58	5	497	87.32%	11.67%	1.01%
G071_BRAZILIAN	507	145	20	672	75.45%	21.58%	2.98%
G072_CHILEAN	1,762	141	10	1,912	92.15%	7.37%	0.52%
G073_ECUADORIAN	255	25	0	280	91.07%	8.93%	0.00%
G074_MEXICAN	1,369	218	22	1,610	85.03%	13.54%	1.37%
G075_OT_L_C_S_AM	3,355	486	63	3,905	85.92%	12.45%	1.61%
G076_PERUVIAN	722	70	13	805	89.69%	8.70%	1.61%
H020_CUBAN	173	47	6	226	76.55%	20.80%	2.65%
H021_HAITIAN	3,135	45	9	3,191	98.25%	1.41%	0.28%
H022_JAMAICAN	3,511	172	11	3,697	94.97%	4.65%	0.30%
H023_OT_CARL_NIE	364	54	18	436	83.49%	12.39%	4.13%
H024_PUERTO_RICO	185	29	2	216	85.65%	13.43%	0.93%
H025_W_IND_NIE	6,901	637	70	7,607	90.72%	8.37%	0.92%
INVALID	4,279	310	137	4,734	90.39%	6.55%	2.89%
I027_W_IND_BLACK	65	1	6	72	90.28%	1.39%	8.33%
I028_BLACK_AMER	15	2	0	17	88.24%	11.78%	0.00%
I029_CAN_BLACK	94	10	1	105	89.52%	9.52%	0.95%
I030_BLACK_NIE	61	3	4	68	89.71%	4.41%	5.88%
I031_AFRI_BLACK	1,254	139	31	1,427	87.88%	9.74%	2.17%
J032_OT_AFRI_NIE	1,968	244	29	2,246	87.62%	10.86%	1.29%
K033_BELGIAN	13,718	1,012	93	14,817	92.58%	6.83%	0.63%
K035_LUXEMBOURG	293	63	9	365	80.27%	17.26%	2.47%
K036_FINNISH	16,602	1,418	125	18,147	91.49%	7.81%	0.69%
K037_DANISH	21,386	1,920	196	23,500	91.00%	8.17%	0.83%
K038_ICELANDIC	10,406	835	88	11,334	91.81%	7.37%	0.78%
K039_NORWEGIAN	42,867	5,392	494	48,755	87.92%	11.06%	1.01%
K040_SWEDISH	35,179	5,404	449	41,039	85.72%	13.17%	1.09%
K041_SCANDI_NIE	6,016	341	40	6,393	94.10%	5.33%	0.63%
L042_AUSTRIAN	13,116	1,674	217	15,207	86.25%	12.32%	1.43%
L043_CZECH	7,005	657	75	7,737	90.54%	8.49%	0.97%
L044_CZECHSLOV	7,566	841	88	8,495	89.08%	9.90%	1.04%
L045_SLOVAK	4,984	329	26	5,339	93.35%	6.16%	0.49%
L046_HUNGARIAN_M	34,237	2,096	193	36,532	93.72%	5.74%	0.53%
L049_SWISS	11,382	863	84	12,330	92.31%	7.00%	0.68%
M050_ALBANIAN	236	31	2	269	87.73%	11.52%	0.74%
M051_BULGAR	1,051	126	14	1,191	88.25%	10.58%	1.18%
M052_CROATIAN	8,053	345	35	8,432	95.51%	4.09%	0.42%
M053_SERBIAN	2,230	210	14	2,454	90.87%	8.56%	0.57%
M054_SLOVENIAN	1,408	145	8	1,559	90.19%	9.30%	0.51%
M055_YUGOSL_NIE	15,067	1,140	94	16,304	92.41%	6.99%	0.58%
M056_MACEDONIAN	2,822	446	27	3,296	85.62%	13.53%	0.82%
M057_GREEK	32,866	666	145	33,677	97.59%	1.98%	0.43%

Table 1: Distribution of Ethnic Origins, ETO401, ETO411, ETO421, 1986 Census Data, Unweighted, Unrounded, E&amp;I, 2B Merge

ORIGIN	ETO401	ETO411	ETO421	TOTAL WRITE-INS	% ETO401 OF TOTAL	% ETO411 OF TOTAL	% ETO421 OF TOTAL
M059_MALTESE	4,475	119	13	4,607	97.13%	2.58%	0.28%
M060_PORTUGUESE	43,839	861	109	44,808	97.84%	1.92%	0.24%
M061_SPANISH	19,033	2,221	266	21,520	88.44%	10.32%	1.24%
M062_OT_EURO_NIE	1,140	203	40	1,383	82.43%	14.68%	2.89%
N063_ESTONIAN	3,633	262	29	3,924	92.58%	6.68%	0.74%
N064_LATVIAN	3,623	248	13	3,884	93.28%	6.39%	0.33%
N065_LITHUANIAN	4,864	308	39	5,213	93.31%	5.91%	0.75%
N066_BYELORUS	350	30	6	386	90.67%	7.77%	1.55%
N067_ROMANIAN	8,756	1,149	115	10,021	87.38%	11.47%	1.15%
N068_RUSSIAN	18,000	1,949	255	20,212	89.06%	9.84%	1.26%
P077_EGYPTIAN	2,739	168	32	2,939	93.19%	5.72%	1.09%
P078_LEBANESE	8,026	448	45	8,520	94.20%	5.26%	0.53%
P079_PALESTINIAN	260	32	7	299	86.96%	10.70%	2.34%
P080_SYRIAN	1,173	196	23	1,392	84.27%	14.08%	1.65%
P081_ARAB_NIE	6,025	453	73	6,552	91.96%	6.91%	1.11%
P082_ISRAELI	181	31	8	219	82.65%	14.16%	3.65%
P083_IRANIAN	2,819	124	26	2,967	95.01%	4.18%	0.88%
P084_TURK	1,270	154	19	1,442	88.07%	10.68%	1.32%
P085_ARMENIAN	4,938	198	20	5,154	95.77%	3.84%	0.39%
Q086_BENGALI	99	12	5	116	85.34%	10.34%	4.31%
Q087_GUJARATI	173	54	5	232	74.57%	23.28%	2.16%
Q088_PUNJABI	1,488	297	8	1,793	82.99%	16.56%	0.45%
Q089_TAMIL	295	119	3	417	70.74%	28.54%	0.72%
Q090_E_INDIA_NIE	46,403	1,723	204	48,330	96.01%	3.57%	0.42%
Q091_BANGLAD_NIE	287	20	2	309	92.88%	6.47%	0.65%
Q092_PAKIST_NIE	5,428	512	42	5,985	90.69%	8.55%	0.70%
Q093_SRI_LAN_NIE	1,292	59	7	1,358	95.14%	4.34%	0.52%
Q094_SINGHALESE	195	11	2	208	93.75%	5.29%	0.96%
R095_BURMESE¹	220	27	4	251	87.65%	10.76%	1.59%
R096_CAMBODIAN	2,133	56	1	2,190	97.40%	2.56%	0.05%
R097_LAOTIAN	2,065	46	0	2,111	97.82%	2.18%	0.00%
R098_THAI	519	50	6	575	90.26%	8.70%	1.04%
R099_VIETNAMESE	11,652	118	23	11,792	98.81%	1.00%	0.20%
R101_KOREAN	5,618	34	5	5,657	99.31%	0.60%	0.09%
R102_JAPANESE	10,271	249	34	10,557	97.29%	2.36%	0.32%
R103_MALAY	395	63	23	481	82.12%	13.10%	4.78%
R104_OT_ASIA_NIE	513	54	19	586	87.54%	9.22%	3.24%
S105_FIJIAN	1,348	155	3	1,506	89.51%	10.29%	0.20%
S106_PHILIPINO	19,835	395	35	20,264	97.88%	1.95%	0.17%
S107_INDONESIAN	647	27	14	688	94.04%	3.92%	2.03%
S108_POLYNESIAN	117	19	4	140	83.57%	13.57%	2.86%
S109_AUST_N_ZEAL	1,457	233	28	1,718	84.81%	13.56%	1.63%
S110_OT_PACIFIJ	271	83	14	368	73.64%	22.55%	3.80%
T111_BUDDHIST	23	1	0	24	95.83%	4.17%	0.00%
T112_HINDU	921	120	10	1,051	87.63%	11.42%	0.95%
T113_SIKH	909	253	33	1,195	76.07%	21.17%	2.76%
T114_MUSLIM_ISL	716	212	37	965	74.20%	21.97%	3.83%
T116_DOUKHOBOR_F	280	58	3	341	82.11%	17.01%	0.88%
T117_HUTTERITE	1,327	51	79	1,457	91.08%	3.50%	5.42%
T118_MENNONITE	4,588	387	25	5,003	91.70%	7.74%	0.50%
T119_OT_RELI_GR	116	25	8	149	77.85%	16.78%	5.37%
U123_GREEK_CYP	84	1	0	85	98.46%	1.54%	0.00%
U124_TURK_CYP	0	0	0	0	---	---	---
U125_CYPRIOTE	86	33	1	120	71.67%	27.50%	0.83%
V555_OTHER_NIE	286	3	2	291	98.28%	1.03%	0.69%

Table 1: Distribution of Ethnic Origins, ETO401, ETO411, ETO421, 1966 Census Data, Unweighted, Unrounded, E&I, 2B Merge

ORIGIN	ETO401	ETO411	ETO421	TOTAL WRITE-INS	% ETO401 OF TOTAL	% ETO411 OF TOTAL	% ETO421 OF TOTAL
W888_NONE	256	18	18	290	88.28%	5.52%	6.21%
Y000_MAR_OF_C_B	8	7	7	22	36.36%	31.82%	31.82%
Z120_EXTR_CODE	0	0	0	0	---	---	---
ALL OTHER	0	0	0	0	---	---	---
TOTAL (B004 TO ALL OTHER)	663,453	51,415	6,030	720,965	92.02%	7.13%	0.84%

## APPENDIX B

### LIST OF STATISTICS CANADA PUBLICATIONS SHOWING 1986 ETHNIC ORIGIN DATA

Catalogue No	Title
93-109	Ethnicity, Immigration and Citizenship
93-156	Census Metropolitan Areas
93-154	Profile of Ethnic Groups
93-155	Profile of the Immigrant Population
98-132	Ethnic Diversity in Canada
99-101E/F	Dictionary
99-104E/F	1986 Census Handbook
98-101 to 98-112	Metropolitan Atlas Series
93-114	Total Income: Individuals
93-118	Family Income: Economic Families
93-153	Language Retention and Transfer
93-157	Canadians and Their Occupations: A Profile
94-101 to 94-124	Census Divisions
94-125, 94-126, 94-133, 94-134	Federal Electoral Districts
94-127 and 94-128	Census Metropolitan Areas and Census Agglomerations
95-101 to 95-174	Census-tracted Centres
94-129 and 94-130	Urban and Rural Areas
94-131	Labour Force Survey Economic Regions
EC86B01	Basic Summary Tables

# APPENDIX C

Population and Occupied Private Dwellings, Showing Estimates for Incompletely Enumerated Indian Reserves and Indian Settlements, for Canada, Provinces and Territories, 1986

	Excluding incom- pletely enumerated Indian reserves and Indian settlements		Estimates for incom- pletely enumerated Indian reserves and Indian settlements		Including estimates on incompletely enumerated Indian reserves and Indian settlements	
	Popula- tion	Occupied private dwellings (1)	Popula- tion	Occupied private dwellings (1)	Popula- tion	Occupied private dwellings (1)
CANADA	25,309,331	9,046,933	44,733	10,600	25,354,064	9,057,533
Newfoundland	568,349	159,917	-	-	568,349	159,917
Prince Edward Island	126,646	40,872	-	-	126,646	40,872
Nova Scotia	873,176	297,224	23	6	873,199	297,230
New Brunswick	709,442	232,701	980	271	710,422	232,972
Quebec	6,532,461	2,370,889	7,815	1,582	6,540,276	2,372,471
Ontario	9,101,694	3,243,752	11,821	3,112	9,113,515	3,246,864
Manitoba	1,063,016	384,324	8,216	1,673	1,071,232	385,997
Saskatchewan	1,009,613	360,467	585	123	1,010,198	360,590
Alberta	2,365,825	840,527	9,453	1,949	2,375,278	842,476
British Columbia	2,883,367	1,094,217	5,840	1,884	2,889,207	1,096,101
Yukon Territory	23,504	8,143	-	-	23,504	8,143
Northwest Territories	52,238	13,900	-	-	52,238	13,900

- (1) Includes all private dwellings occupied by usual residents as well as private dwellings occupied solely by foreign and/or temporary residents.

**APPENDIX D**  
**MOTHER TONGUE AND ETHNIC ORIGIN CODING ERRORS**

Miscodes Between Mother Tongue and Ethnic Origin, Canada, 1986

Mother Tongue	Ethnic Origin	Number of Cases
Armenian	Luxembourg	45
Czech	Puerto Rican	100
Serbian	Punjabi	65
Croatian	East Indian, NIE	85
Arabic	Peruvian	320
Persian	Icelandic	60
Japanese	Swiss	75
Khmer	Iranian	65
Icelandic	Acadian	20
Bengali	Norwegian	100
Punjabi	Finnish	40
Urdu	Czech	35
Singhalese	Swedish	75
Korean	Albanian	25
Chillock	Russian	50
Micmac	Yugoslav, NIE	90
Slovak	Other West Indian	85
Yugoslavian, NIE	Other Caribbean, NIE	100
Latvian	East Indian, NIE	75
Ojibway	Croatian	25

1986 Code and Equivalent Entries for Mother Tongue, Home Language and Ethnic Origin

<u>CODE</u>	<u>MOTHER TONGUE</u>	<u>HOME LANGUAGE</u>	<u>ETHNIC ORIGIN</u>
001	Portuguese	Portuguese	
002	Spanish	Spanish	
003	Romanian	Romanian	
004	Dutch	Dutch	British N.I.E.
005	Flemish	Flemish	Other British
006	Frisian	Frisian	Welsh
007	Danish	Danish	
008	Icelandic	Icelandic	Acadian
009	Norwegian	Norwegian	Franco-Manitoban
010	Swedish	Swedish	Franco-Ontarian
011	Yiddish	Yiddish	Québécois
012	Alsacian	Alsacian	French Canadian
013	Gaelic	Gaelic	
014	Welsh	Welsh	
015	Irish	Irish	
016	Scottish	Scottish	Amerindian N.I.E.
017	Other Celtic	Other Celtic	Canadian
018	Russian	Russian	
019	Byelorussian	Byelorussian	American
020	Bulgarian	Bulgarian	
021	Serbo Croate	Serbo Croate	Haitian
022	Slovene	Slovene	Jamaican
023	Yugoslav N.I.E.	Yugoslav N.I.E.	Other Caribbean
024	Czech	Czech	Puerto Rican



<u>CODE</u>	<u>MOTHER TONGUE</u>	<u>HOME LANGUAGE</u>	<u>ETHNIC ORIGIN</u>
025	Slovak	Slovak	West Indian N.I.E. French *
026	Macedonian	Macedonian	English *
027	English *	Polish	West Indian Black Irish *
028	French *	Other Slavic	Black American Scottish *
029	Italian *	Greek	Canadian Black German *
030	German *	Armenian	Black N.I.E. Italian *
031	Ukrainian *	Lettish	African Black Ukrainian *
032	Polish	Lithuanian	Other African N.I.E. Dutch *
033	Other Slavic Languages	Iranian	Belgian Chinese *
034	Greek	Bengali	Jewish *
035	Armenian	Cingalese	Luxembourg Polish *
036	Lettish	Hindi	Finnish Black *
037	Lithuanian	Punjabi	Danish Inuit *
038	Iranian	Urdu	Icelandic North American Indian *
039	Bengali	Other Indo Pakistani	Norwegian Métis *
040	Cingalese	Turkish Languages	Swedish
041	Hindi	Estonian	Scandinavian N.I.E.
042	Punjabi	Finnish	Austrian

<u>CODE</u>	<u>MOTHER TONGUE</u>	<u>HOME LANGUAGE</u>	<u>ETHNIC ORIGIN</u>
043	Urdu	English *	Czech
044	Other Indo-Pakistani	French *	Czechoslovakian
045	Turkish Languages	Italian *	Slovak
046	Estonian	Chinese *	-
047	Finnish	German *	-
048	Hungarian	Hungarian	Hungarian
049	Japanese	Japanese	Swiss
050	Korean	Korean	Albanian
051	Cree	Cree	Bulgarian
052	Ojibway	Ojibway	Croatian
053	Black foot	Black foot	Serbian
054	Malecite	Malecite	Slovenian
055	Micmac	Micmac	Yugosi, N.I.E.
056	Montagnais Naskapi	Montagnais Naskapi	Macedonian
057	Other Algonquin	Other Algonquin	Greek
058	Chipewyan	Chipewyan	-
059	Slave	Slave	Maltese
060	Carrier	Carrier	Portugese
061	Other Athapaskan	Other Athapaskan	Spanish
062	Haida	Haida	Other European N.I.E.
063	Iroquois	Iroquois	Estonian
064	Kutenai	Kutenai	Latvian
065	Salishan	Salishan	Lithuanian
066	Sioux (Dakota)	Sioux (Dakota)	Byelorussian
067	Tlingit	Tlingit	-

<u>CODE</u>	<u>MOTHER TONGUE</u>	<u>HOME LANGUAGE</u>	<u>ETHNIC ORIGIN</u>
068	Chillock	Chillock	Russian
069	Tsimshian	Tsimshian	-
070	Wakashan	Wakashan	Argentinian
071	Other Amerindian	Other Amerindian	Brazilian
072	Inuktitut	Inuktitut	Chilean
073	Malayalam	Malayalam	Ecuadorian
074	Tamil	Tamil	Mexican
075	Telugu	Telugu	Other Latin/Central/South
076	Arabic	Arabic	Peruvian
077	Hebrew	Hebrew	Egyptian
078	Maltese	Maltese	Lebanese
079	Other Semitic	Other Semitic	Palestinian
080	Thai	Thai	Syrian
081	Vietnamese	Vietnamese	Arab N.I.E.
082	Other Asian	Other Asian	Israeli
083	Khmer	Khmer	Iranian
084	Malay	Malay	Turk
085	Pilipino	Pilipino	Armenian
086	Tagalog	Tagalog	Bengali
087	Other Malayo-Polynesian	Other Malayo-Polynesian	Gujarati
088	Swahili	Swahili	Punjabi
089	Other Bantu	Other Bantu	Tamil
090	Other Nigero-Congolese	Other Nigero-Congolese	East Indian N.I.E.
091	Other African	Other African	Bangladeshi N.I.E.
094	Chinese	-	Singhalese

<u>CODE</u>	<u>MOTHER TONGUE</u>	<u>HOME LANGUAGE</u>	<u>ETHNIC ORIGIN</u>
096	Other	Other	Cambodian
097	Creole	Creole	Laotian
098	Belgian	Blegian	Thai
100	Mohawk	Mohawk	-
101	Luxembourg	Luxembourg	Korean
102	Scandinavian	Scandinavian	Japanese
103	Swiss	Swiss	Malay
104	Kutchin	Kutchin	Other Asian N.I.E.
105	Hare	Hare	Fijian
106	Dogrib	Dogrib	Filipino
107	Yellowknife	Yellowknife	Indonesian
108	Nahani	Nahani	Polynesian
109	Tahltan	Tahltan	Australian/New Zealander
110	Chilcotin	Chilcotin	Other Pacific Islanders
111	Tutchone	Tutchone	Buddhist

\* Corresponds to mark-in box on questionnaire

[illegible]

